# Correction of sensor crosstalk error in Exhalyzer D multiple-breath washout device significantly impacts outcomes in children with cystic fibrosis

Florian Wyler[†1] (https://orcid.org/0000-0002-1232-1392)

Marc-Alexander H. Oestreich[†1,2] (https://orcid.org/0000-0001-9641-3691)

Bettina S. Frauchiger[1,3] (https://orcid.org/0000-0002-9519-9328)

Kathryn A. Ramsey[#1] (https://orcid.org/0000-0003-4574-6917)

Philipp T. Latzin[*#1] (https://orcid.org/0000-0002-5239-1571)

[†] Authors contributed equally to this work as first author

[#] Authors contributed equally to this work as last author

* Corresponding author

[1] Division of Paediatric Respiratory Medicine and Allergology, Department of Paediatrics, Inselspital, Bern University Hospital, University of Bern, Switzerland

[2] Graduate School for Health Sciences, University of Bern, Switzerland.

[3] Graduate School of Cellular and Biomedical Sciences, University of Bern, Switzerland

**Author contributions:**

F.W., M.O., K.R. and P.L contributed to the design of the study; F.W. and M.O. collected and analyzed the data. M.O., B.F and K.R created and organized the databases for retrospective analysis. F.W. drafted the manuscript, with major contributions from M.O, K.R. and P.L. All authors approved the final version of the manuscript.

**Running title:** Measurement error in a multiple-breath washout device

**Corresponding Author:**

Prof. Philipp Latzin, MD, PhD

Inselspital, University Children's Hospital of Bern

Freiburgstrasse 15, 3010 Bern, Switzerland

E-mail: philipp.latzin@insel.ch

Phone: +41 31 632 9353

**Total word count:** 393 Intro + 492 Methods + 780 Results + 1826 Discussion **= 3491**

**Online Supplement:** This manuscript has an online supplement, accessible here:

DOI: https://doi.org/10.6084/m9.figshare.14604846

URL: https://figshare.com/articles/online_resource/_/14604846

**Presentations**: Preliminary results were presented at the ERS international congress 2020 (**DOI:** 10.1183/13993003.congress-2020.1171).

## ABSTRACT

Rationale: Nitrogen multiple-breath washout is an established technique to assess functional residual capacity and ventilation inhomogeneity in the lung. Accurate measurement of gas concentrations is essential for the appropriate calculation of clinical outcomes.

Objectives: We investigated the accuracy of oxygen and carbon dioxide gas sensor measurements used for the indirect calculation of nitrogen concentration in a commercial multiple-breath washout device (Exhalyzer D, Eco Medics AG, Duernten, Switzerland) and its impact on functional residual capacity and lung clearance index.

Methods: High precision calibration gas mixtures and mass spectrometry were used to evaluate sensor output. We assessed the impact of corrected signal processing on multiple-breath washout outcomes in a dataset of healthy children and children with cystic fibrosis using custom analysis software.

Results: We found inadequate correction for the cross sensitivity of the oxygen and carbon dioxide sensors in the Exhalyzer D device. This results in an overestimation of expired nitrogen concentration, and consequently multiple-breath washout outcomes. Breath-by-breath correction of this error reduced the mean (SD) cumulative expired volume by 19.6 (5.0)%, functional residual capacity by 8.9 (2.2)%, and lung clearance index by 11.9 (4.0)%. It also substantially reduced the level of the tissue nitrogen signal at the end of measurements.

Conclusions: Inadequate correction for cross sensitivity in the oxygen and carbon dioxide gas sensors of the Exhalyzer D device leads to an overestimation of functional residual capacity and lung clearance index. Correction of this error is

possible and could be applied by re-analyzing the measurements in an updated software version.

## NEW AND NOTEWORTHY

We investigated the sensor accuracy of a prominent nitrogen multiple-breath washout ($N_2$MBW) device (Eco Medics AG, Duernten, Switzerland) as a possible cause of lack of comparability between outcomes of different MBW devices and methods. We identified an error in the nitrogen concentration calculation of this device, which results in a 10-15% overestimation of primary outcomes, functional residual capacity and lung clearance index. It also leads to a significant overestimation of nitrogen back-diffusion into the lungs.

## INTRODUCTION

Cystic fibrosis (CF) is a chronic, genetic disease characterized by lung function decline and respiratory failure. Newborn screening and early interventions for CF have resulted in the majority children with CF having no overt respiratory symptoms and normal spirometry(1). However, structural lung disease is present early on high-resolution chest CT scans and progresses during childhood(2). Therefore, sensitive functional outcomes are needed to monitor disease progression and assess treatment responses in children with CF.

The multiple breath washout (MBW) technique is more sensitive than spirometry to detect early CF lung disease(3, 4). The lung clearance index (LCI) from MBW correlates with underlying structural lung disease and tracks disease progression in children with CF(1, 5-7). LCI has been endorsed as an endpoint in clinical trials in children and adults with CF in North American and Europe(8-10). In young children

with CF, LCI significantly improved in response to disease modifying therapies and hypertonic saline, while spirometry outcomes did not change(11, 12). Therefore, the LCI is a promising endpoint for clinical management and interventional trials in children with CF, even more in the era of CFTR modulator therapies.

The most common MBW technique used in clinical trials is the nitrogen washout ($N_2$MBW), using the Exhalyzer D device (Eco Medics AG, Duernten, Switzerland)(12, 13). The subject breathes 100% oxygen to wash out resident nitrogen to 2.5% of the starting concentration. However, this technique relies on indirect calculation of $N_2$ concentration through the measurement of oxygen ($O_2$) and carbon dioxide ($CO_2$), which means that small errors in the $O_2$ or $CO_2$ concentration could result in instantaneous and cumulative errors in the $N_2$ concentration, particularly at the end of test criteria(14). Further, as $N_2$ is soluble in the blood and tissues, there is the potential for $N_2$ diffusion from the lung tissue into the alveolar spaces during the washout(15). A combination of these effects has been proposed to be the cause of elevated $N_2$MBW outcomes as compared to $SF_6$MBW(16). It is essential to address these concerns in $N_2$ methodology to ensure the appropriate calculation of clinical outcomes.

In this study we investigated the accuracy of indirect nitrogen measurement using the Exhalyzer D device by i) assessing the sensor accuracy of its $O_2$ and $CO_2$ sensors, ii) establishing a correction for any observed sensor error and iii) assessing the effect size of the sensor error on clinical outcomes and tissue nitrogen.

## METHODS

### Study design

This was an experimental study to assess the sensor accuracy of the Exhalyzer D device. We also performed a retrospective analysis of existing $N_2$MBW data to assess the impact of sensor inaccuracy on MBW outcomes. The Ethics Committee of the Canton of Bern, Switzerland approved the study protocol (PB_2017-02139).

### i) Sensor accuracy

To assess sensor accuracy over the wide range of concentrations encountered in a $N_2$MBW measurement, we collected experimental data from gas mixture measurements and mass spectrometry measurements. We compared the measured $O_2$ and $CO_2$ concentrations to the known gas mixture concentrations.

#### Gas mixtures

Sensor accuracy was assessed over a representative range of concentrations present in $N_2$MBW measurements. Twelve technical gas mixtures (Carbagas AG, Muri bei Bern, Switzerland) were used, each containing different combinations of $CO_2$, $O_2$ and $N_2$ concentrations.

Additionally, a series of mass spectrometry measurements was carried out, where $N_2$ was kept at 2% to mimic the MBW end of test condition, while $CO_2$ and $O_2$ were varied (AMIS 2000 Mass Spectrometer, Innovision ApS, Odense, Denmark). Mass spectrometry data was provided by Eco Medics AG, Duernten, Switzerland.

#### Sensor characteristics

The Exhalyzer D measures both $O_2$ (X3004 OEM sensor, Oxigraf Inc., Sunnyvale, CA, USA) and $CO_2$ (Capnostat 5, Respironics Inc., Wallingford, CT, USA) using

absorption spectroscopy, a technique where the absorption of light is measured at specific frequencies that are characteristic to each gas(17). The absorption spectra of $O_2$ and $CO_2$ are affected by a variety of factors, including pressure, temperature, and the presence of other gases(18, 19). This leads to a sensor cross sensitivity, where the presence of $O_2$ in the gas mixture can affect the absorption spectrum (and therefore measured concentration) of $CO_2$ and vice versa.

**ii) Correction function**

We combined the data of the technical gas mixture and mass spectrometry measurements to construct a correction function for the $O_2$ and $CO_2$ sensors. We fitted a $2^{nd}$-degree two-parameter polynomial through the error for each sensor, as a function of measured $O_2$ and $CO_2$ (Eq. E4 in the online supplement). Fitting was performed using MATLAB 2017b (Mathworks, Natick, Massachusetts, USA). This characterization of the measurement error as a function of measured gas concentrations could then be directly used as a correction function for the analysis of MBW measurements. For each combination of $O_2$ and $CO_2$ we added the fitted error for each sensor to the respective measured concentrations. Note that when characterizing the error currently present in the sensors, the existing linear $CO_2$ crosstalk correction of Spiroware 3.2.1 was still applied(20), but when characterizing the correction function, parameters were chosen to replace and improve the existing crosstalk correction.

**iii) Impact on outcomes**

We characterized the impact of measurement error on MBW outcomes in 357 measurements from 85 healthy children(21) and 62 children diagnosed with CF(5, 22) (Table E4 in the online supplement). We compared outcomes between those

158 calculated using standard analysis algorithms (replicating Spiroware 3.3) and
159 corrected algorithms using a custom Python script developed by our group.

160

161

162 **RESULTS**

163 **i) Sensor accuracy**

164 We found that the Exhalyzer D device has $O_2$ and $CO_2$ sensor measurement errors
165 which result in overestimated $N_2$ concentrations. The $N_2$ error was non-linear and $O_2$-
166 and $CO_2$-dependent, with the highest error occurring at very high $O_2$ concentrations
167 and increasing $CO_2$ concentrations (Figure 1). Therefore, at the standard end of test
168 conditions ($1/40^{th}$ starting $N_2$ concentration: end expiratory $N_2$ at 2%, $CO_2$ at 5%, $O_2$
169 at 93%), the $N_2$ concentration was significantly overestimated (Table 1). At the
170 original end of test, the Exhalyzer D measures 2% $N_2$ when the real $N_2$ concentration
171 was 1.1%. Correction for the sensor error reduces $N_2$ concentrations at the end of the
172 washout and the end of test criteria was systematically reached earlier (Figure 2). At
173 the new end of test conditions, the relative error in $N_2$ concentration following
174 correction was estimated to be 44% (2.88% $N_2$ standard vs 2% $N_2$ corrected; Table
175 1).

**Table 1**

| Condition | Signal Processing | | | | | | $N_2$ Error | | Contribution | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Standard | | | Corrected | | | | | | |
| | $[N_2]$ | $[CO_2]$ | $[O_2]$ | $[N_2]$ | $[CO_2]$ | $[O_2]$ | abs | rel | $[CO_2]$ | $[O_2]$ |
| Original end of test [%] | 2.00 | 5.00 | 93.0 | 1.10 | 5.12 | 93.8 | 0.90 | 82.4 | 13.5 | 86.5 |
| New end of test [%] | 2.88 | 4.88 | 92.2 | 2.00 | 5.00 | 93.0 | 0.88 | 44.1 | 13.3 | 86.7 |
| No nitrogen [%] | 0.88 | 4.88 | 94.2 | 0.00 | 5.00 | 95.0 | 0.88 | - | - | - |

**Table 1**: Specific examples of sensor impact on measurement of $N_2$ in three conditions of interest. The original end of test corresponds to a gas mixture that would be identified as the end of test in standard processing. The second condition corresponds to the new end of test after sensor correction. The third condition contains no real nitrogen. Standard concentrations denote concentrations measured in standard Spiroware 3.3 processing. Corrected concentrations correspond to concentrations after sensor correction is applied. $N_2$ error summarizes the absolute (abs) difference between $N_2$ in standard vs. corrected, as well as the relative (rel) error ((standard-corrected)/corrected). The relative contribution of each sensor in [%] to the total error in $N_2$ concentration is listed under “Contribution”.

$O_2$ sensor

The majority of the error in $N_2$ measurement (87% of the error at the test end, see Table 1) originated from the $O_2$ sensor of the Exhalyzer D device. The $O_2$ measurement error was non-linear and dependent on the $CO_2$ concentration (Figure 3A). This error resulted in underestimation of $O_2$ concentrations, with a greater underestimation with increasing concentrations of $CO_2$ in the measured range of 0 – 7.5%. This in turn lead to an overestimation of calculated concentrations of $N_2$ as described above.

$CO_2$ sensor

We also found an error in the $CO_2$ concentration. The $CO_2$ sensor output was already corrected by a factor that depends on the concentration of $O_2$ (Eq. E3 in the online supplement). However, the $CO_2$ sensor seemed to display a different cross-sensitivity than the standard signal processing takes into account (Figure 3B). The residual error appeared to be primarily dependent on $CO_2$ concentration and only partially dependent on $O_2$ concentration. The $CO_2$ error also lead to an overestimation of $N_2$, however the impact of the $CO_2$ error was smaller than the $O_2$ error, making up 13% of the total sensor error at the end of test conditions (Table 1).

**iii) Effect size of sensor correction**

Sensor correction impact on MBW outcomes

We re-analyzed 357 MBW measurements from healthy controls (HC) and children with CF using the sensor correction functions outlined above in a custom software. Application of the $O_2$ and $CO_2$ sensor correction functions had a significant impact on all MBW outcomes (Table 1). Following the sensor correction, the mean (SD) cumulative expired volume decreased by 19.6 (5.0) %, FRC decreased by 8.9 (2.2)

%, and LCI decreased by 11.9 (4.0) %. The reduced CEV is explained by lower end-expiratory concentrations of $N_2$, which lead to an earlier end of test (i.e. the criterion of reaching $1/40^{th}$ of the initial $N_2$ concentration is reached earlier). Decreased FRC is explained by slightly lower concentrations of $N_2$ throughout the measurement. The decrease in CEV was more pronounced than for FRC, and with LCI being the ratio of those two outcomes (LCI = CEV/FRC), this leads to an overall decrease in LCI. The change in outcomes following sensor correction could vary greatly for individual measurements (Table 2). However, outcomes before and after the correction over a large number of measurements correlate strongly. Linear fits of corrected outcomes vs standard outcomes have $R^2$ values of 0.997 for FRC, and 0.96 for LCI (Figure 4), respectively.

The significance of differences in LCI and FRC [L] observed between healthy controls and children with CF present in the uncorrected data were also present following sensor correction (Table 2). The change in outcomes following correction was dependent on the magnitude of the outcomes themselves for both FRC and LCI (Figure 5, and OLS Figure 1).

225 **Table 2**

226

| | | | Standard | | | Corrected | | | Difference | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | n | mean | SD | p-value* | mean | SD | p-value* | mean | rel [%] | 95% CI [%] | | | p-value† |
| **LCI** [TO] | All | 147 | 8.33 | 2.05 | | 7.31 | 1.7 | | 1.02 | 11.9 | 11.2 | - | 12.5 | **<0.001** |
| | HC | 85 | 7.12 | 0.51 | | 6.30 | 0.4 | | 0.82 | 11.3 | 10.6 | - | 12.1 | **<0.001** |
| | CF | 62 | 9.99 | 2.21 | | 8.69 | 1.8 | | 1.30 | 12.6 | 11.4 | - | 13.8 | **<0.001** |
| | Difference | | -2.87 | | **<0.001** | -2.38 | | **<0.001** | | | | | | |
| **FRC** [L] | All | 147 | 1.63 | 0.87 | | 1.49 | 0.80 | | 0.14 | 8.9 | 8.6 | - | 9.3 | **<0.001** |
| | HC | 85 | 1.87 | 0.95 | | 1.73 | 0.89 | | 0.14 | 7.9 | 7.6 | - | 8.1 | **<0.001** |
| | CF | 62 | 1.31 | 0.61 | | 1.17 | 0.53 | | 0.14 | 10.4 | 9.7 | - | 11.0 | **<0.001** |
| | Difference | | 0.56 | | **<0.001** | 0.56 | | **<0.001** | | | | | | |
| **CEV** [L] | All | 147 | 14.9 | 7.4 | | 11.9 | 5.7 | | 3.03 | 19.6 | 18.8 | - | 20.4 | **<0.001** |
| | HC | 85 | 14.6 | 6.7 | | 12.0 | 5.6 | | 2.67 | 18.2 | 17.4 | - | 19.0 | **<0.001** |
| | CF | 62 | 15.2 | 8.3 | | 11.7 | 5.8 | | 3.54 | 21.5 | 20.1 | - | 23.0 | **<0.001** |
| | Difference | | -0.58 | | 0.6532 | 0.29 | | 0.7601 | | | | | | |

227

228 **Table 2**: Summary of the differences in Lung Clearance Index (LCI), functional residual capacity (FRC) and cumulative expired volume (CEV) between
229 healthy controls (HC) and and patients with cystic fibrosis (CF) in the retrospective dataset before (standard) and after (corrected) the application of the sensor
230 correction function. *unpaired t test; †paired t test. Bold print indicates statistical significance.

### Sensor correction impact on tissue nitrogen

We also observed a substantial impact of the sensor corrections on tissue nitrogen. Towards the end of a MBW measurement, the concentration of $N_2$ in the lung drops so low that diffusion of $N_2$ from the body becomes a potential concern for the accuracy of the MBW outcomes. In the Exhalyzer D, $N_2$ concentration is currently overestimated in the presence of $CO_2$ (i.e. during expirations), even in the complete absence of $N_2$ (Figure 6B). In conditions reflecting expirations where there is no $N_2$ exhaled ($CO_2$ around 5%, rest $O_2$), the Exhalyzer D still measures a concentration of $N_2$ of 0.88% (Table 1, and Figure 6B, intersection of 5% line with x-axis). As correction of the $O_2$ and $CO_2$ error significantly reduces the $N_2$ concentration, a significant part of the tissue nitrogen signal at a diffusion equilibrium in long MBW measurements disappears after correction (Figure 6A). The higher the end-expiratory concentrations of $CO_2$, the greater this effect (Figure 6B).

## DISCUSSION

### Summary

We report a significant measurement error in the Eco Medics Exhalyzer D $N_2$MBW device. At high concentrations of $O_2$, and natural end-expiratory concentrations of $CO_2$, the device's sensors underestimate $O_2$ and $CO_2$ gas concentrations and it therefore overestimates end-expiratory concentrations of $N_2$. Artificial elevation of $N_2$ during the washout influences the end of test criterion and causes overestimation of FRC and LCI. It also results in a significant overestimation of measured tissue nitrogen at the end of the test.

**i) Sensor accuracy**

We are the first group to formally characterize this sensor cross sensitivity error. Previous studies have reported high expiratory $N_2$ concentrations at the end of long MBW measurements(23-25). While this has predominantly been attributed to the release of $N_2$ from the lung tissue(15), others have argued that tissue nitrogen alone may not be sufficient to explain the observed concentrations of $N_2$, and that there may be an additional "offset error" present, speculated to be caused by $CO_2$-crosstalk with the $O_2$-sensor(16). We confirm this impact of sensor crosstalk and comprehensively characterize a significant sensor error in the Exhalyzer D device which is primarily responsible for elevated expiratory $N_2$ concentrations in $N_2$MBW measurements.

Previous validation studies using *in vitro* lung models as well as the internal testing of Eco Medics AG either did not specifically examine the end of test, end-expiratory conditions examined here, or potentially washed $CO_2$ out of the validation system before the end of test condition could be reached(26). This may have made it difficult to identify the impact of sensor cross sensitivity in the critical end of test phase of MBW. It is worth noting that individual sensor errors were relatively small (~1% relative error of a sensor reading), even in the most extreme case (low $N_2$, high $CO_2$). However, the indirect calculation of $N_2$ by the Exhalyzer D device is vulnerable to errors in the high $O_2$ and high $CO_2$ concentrations that occur at the end of the MBW measurement(14), leading to a relative $N_2$ error in this condition of 44%. This measurement error exceeds the recommendations for manufacturers outlined in the ATS/ERS consensus statement of measuring tracer gas concentration within 5% accuracy(27).

The measurements performed here highlight the need for more robust methods of validation for MBW devices. Ideally, such a validation would involve an in vitro lung model that can realistically reproduce the signal dynamics introduced by breathing, and allow for direct comparison of outcomes between devices. In the absence of such a validation system, individual components of MBW devices such as the measurement of tracer gas should be further validated, covering the entire range of concentrations encountered in a MBW measurement. Testing MBW equipment with a wider range of technical gas mixtures with a special emphasis on the test end criterion is highly feasible, and should be a minimum requirement for equipment validation. The specific error in tracer gas measurement described here has been shown to be highly relevant to the Exhalyzer D, but any device that relies on indirect assessment of tracer gas concentrations is potentially vulnerable to cumulative errors in their individual gas sensors.

**ii) Correction function**

The sensor error observed in this study appears systematic and reproducible across Exhalyzer D devices. The correction function required to correct for the sensor error is simple and has now been implemented in the signal processing of Spiroware (3.3.1), and can also be applied retrospectively to existing data. Sandvik et al. accessed the correction factors and the equations from Eco Medics AG, which have also been published by us in preprint form(28). They applied the equations in a custom-made software version to measurements of healthy infants and toddlers(20). They were able to show that after application of these equations, agreement between $N_2$-MBW and $SF_6$-MBW outcomes was closer than without the correction. This and our work suggest that the same equations can be applied to infants and school age children. Notably the correction function suggested here would replace the currently

existing $CO_2$ sensor crosstalk correction. The chosen degree of the polynomial fit constitutes an empirical correction and is a compromise between improving the currently either missing or linear empirical correction and the limits of precision imposed by the intrinsic uncertainty of the reference gas mixtures. To ensure accurate re-analysis of outcomes, this correction function will need to be applied to raw signals on a breath-by-breath basis.

**iii) Effect size of sensor correction**

Sensor correction impact on MBW outcomes

The sensor error described here leads to substantially inflated MBW outcomes. This result provides a new perspective on previously described differences between $N_2$ and $SF_6$ MBW measurements(23-25). It also offers a potential explanation for the differences observed between $N_2$MBW outcomes measured using the Exhalyzer D and devices by other manufacturers such as ndd Medizintechnik AG (Zürich, Switzerland)(29). The primary $N_2$MBW outcomes from the Exhalyzer D were consistently higher than $SF_6$MBW outcomes and $N_2$MBW outcomes from the ndd device. These observations may be partly explained by the systematic overestimation of $N_2$ by the Exhalyzer D reported in this study. The direction of the change after correction suggests that differences between devices will now be smaller. The sensor correction described here has since been used by Sandvik et al. to confirm that agreement between $N_2$MBW and $SF_6$MBW improves upon correction in infants and toddlers(20). In order to validate this in detail, original data need to be reloaded using the sensor correction described here. Fortunately, the $N_2$ error has been an overestimation rather than an underestimation, as measurements can now be re-analyzed without the worry that the trials might not have recorded data long enough to reach the end of test in the corrected measurement.

329 Notably, using the sensor correction detailed here will lead to substantially shortened
330 measurement times for $N_2$MBW. The observed 19.6% reduction in patient breathing
331 required implies that after the sensor correction, the washout portion of the $N_2$MBW
332 measurements would on average be shortened by almost $1/5^{th}$. Despite the strong
333 feasibility of $N_2$MBW in young children within research studies, challenges have been
334 reported when translating to time-limited busier clinical environments(30). Shorter
335 test duration may improve this. Previous studies have tried to reduce MBW test
336 duration by using an earlier LCI cut-off(31) (LCI 5%) or reducing the number of trials
337 used for outcome reporting(32). The sensor correction described here shortens the
338 $N_2$MBW test length without the need to adjusted test protocols or outcomes. With the
339 correction integrated into the Ecomedics software, the reduced test duration for
340 prospective data collection may therefore help to facilitate the transition of $N_2$MBW
341 into the clinical setting(5, 33). The effect of the correction function on other MBW
342 indices such as those calculated by concentraton normalised phase III slope analysis
343 (SnIII) remains yet unclear and needs to be examined in future studies.

344 A major concern that arises with the publication of this study is that it calls into
345 question previously published results obtained using the Exhalyzer D. As the change
346 in outcomes depends on the breathing pattern and $CO_2$ concentrations, it is difficult
347 for users to predict how much outcomes of a single measurement will change. In
348 addition, the change in outcomes will be higher in children with lung disease and
349 elevated MBW outcomes, compared with healthy children. It is to be expected that
350 effect sizes and confidence intervals of MBW outcomes in such studies will change.
351 This also means that reference values or upper limits of normality generated using
352 the Exhalyzer D device will change(21). Previously collected results from ongoing

353 studies will need to be recalculated in order for them to be interpretable alongside
354 values obtained using this correction.

355 However, while the impact of the sensor error has effects which are difficult to predict
356 on the level of individual measurements, the impact on MBW outcomes on a large
357 enough number of files appears more systematic. Whether or not results from
358 previous studies are affected can only reliably be elucidated by re-analysis of raw
359 data. Notably, the impact of the error on outcomes is dependent on their magnitude,
360 therefore the correction is likely to influence outcomes from individuals with lung
361 disease more than healthy controls. Even during the retrospective re-analysis within
362 this study, we observed a change in significance in FRC differences (when
363 normalized by body weight) between healthy children and children with CF (Table E5
364 in online supplement). In addition, overestimated values of LCI may have influenced
365 individual eligibility to enter clinical trials. Re-analysis of MBW measurements used in
366 clinical trials where drug approval was or is based on affected $N_2$MBW data should
367 be prioritized.

368 Sensor correction impact on tissue nitrogen

369 It has been hypothesized that towards the end of a $N_2$MBW test the concentration of
370 $N_2$ in the lungs drops so low that a noticeable amount of $N_2$ diffuses from the body
371 into the lungs(25, 34). Recent lung modelling work suggests that this $N_2$ diffusion is
372 related to local ventilation/perfusion mismatch(35). The results of this current study
373 suggest that the impact of tissue $N_2$ diffusion is significantly lower than previously
374 estimated. Even if no $N_2$ diffused into the lungs, the Exhalyzer D would still measure
375 end-expiratory ($CO_2$ around 5%) concentrations of $N_2$ of about 0.88%, which would
376 significantly perturb estimates of tissue nitrogen. The sensor correction functions

introduced in this paper would therefore reduce a substantial part of the observed tissue nitrogen in measurements performed with the Exhalyzer D.

### Strengths and limitations

Through detailed understanding of the underlying signal processing of the Exhalyzer D we were able to characterize the precise impact of an observed error in gas sensors on the clinical outcomes LCI and FRC. The findings from the technical gases were confirmed by measurements using a mass spectrometer. Using these data, we were able to estimate the impact of the measurement error and develop an appropriate correction function.

The main limitation of this study is the fact that we only had a finite number of gas samples with finite precisions to test the sensors. We chose a selection of gas concentrations from our range of interest which would exhibit cross-sensitivity effects but could ultimately not cover the entire range of concentration combinations in MBW measurements using technical gases. However, the phase of the measurement where sensor accuracy is the most relevant for accurate MBW outcomes is the end of test, whereby the mass spectrometry measurements allowed us to describe the sensor error with high certainty.

### Outlook

In the process of conducting the research for this paper, we contacted the manufacturer for information regarding their sensor configurations and questions regarding sensor settings and signal processing. The have incorporated the correction described in this manuscript into the signal processing of the new software version of Spiroware 3.3.1, which has since been released.

## Conclusion

An error in the cross sensitivity correction between the oxygen and carbon dioxide gas sensors of the Exhalyzer D device leads to an overestimation of FRC and LCI. Correction of this error is possible but needs to be applied breath-by-breath by re-analyzing the measurements in an updated version of the Spiroware analysis software.

## ACKNOWLEDGMENTS

The authors thank Daniel Oberli and Ruedi Isler (Eco Medics AG, Switzerland) for confirming the sensor error after we detected and informed them about it. We also thank Daniel Oberli and Ruedi Isler for working with us to finalize the correction equation based on the technical gas measurements performed by us and the mass spectrometry measurements performed by them, which have now been implemented into the commercial software. We thank all our patients and study participants as well as their families for allowing their MBW data to be used for research. In addition, we thank the lab technicians Sandra Lüscher, Bettina Vessaz, Sharon Krattinger, and Gisela Wirz for their contribution by performing MBW measurements. Special thanks go to the physicians from our department.

## GRANTS, GIFTS, EQUIPMENT, DRUGS

Eco Medics AG (Duernten, Switzerland) provided a research version of their commercial software Spiroware 3.3, provided information on signal processing algorithms, and helped with the acquisition of mass spectrometry measurements.

This project was funded by the Swiss National Science Foundation, Grant Nr. 182719 (P. Latzin) and 168173 (K. Ramsey).

## REFERENCES


1. **Stanojevic S, Davis SD, Retsch-Bogart G, Webster H, Davis M, Johnson RC, Jensen R, Pizarro ME, Kane M, Clem CC, Schornick L, Subbarao P, and Ratjen FA**. Progression of Lung Disease in Preschool Patients with Cystic Fibrosis. *American journal of respiratory and critical care medicine* 195: 1216-1225, 2017.

2. **Mott LS, Park J, Murray CP, Gangell CL, de Klerk NH, Robinson PJ, Robertson CF, Ranganathan SC, Sly PD, and Stick SM**. Progression of early structural lung disease in young children with cystic fibrosis assessed using CT. *Thorax* 67: 509-516, 2012.

3. **Gustafsson PM, De Jong PA, Tiddens HA, and Lindblad A**. Multiple-breath inert gas washout and spirometry versus structural lung disease in cystic fibrosis. *Thorax* 63: 129-134, 2008.

4. **Owens CM, Aurora P, Stanojevic S, Bush A, Wade A, Oliver C, Calder A, Price J, Carr SB, Shankar A, and Stocks J**. Lung Clearance Index and HRCT are complementary markers of lung abnormalities in young children with CF. *Thorax* 66: 481-488, 2011.

5. **Frauchiger BS, Binggeli S, Yammine S, Spycher B, Krüger L, Ramsey KA, and Latzin P**. Longitudinal Course of Clinical Lung Clearance Index in Children with Cystic Fibrosis. *The European respiratory journal* in press, 2020.

6. **Perrem L, Stanojevic S, Shaw M, Jensen R, McDonald N, Isaac SM, Davis M, Clem C, Guido J, Jara S, France L, Solomon M, Grasemann H, Waters V, Sweezey N, Sanders DB, Davis SD, and Ratjen F**. Lung Clearance Index to Track Acute Respiratory Events in School-Age Children with Cystic Fibrosis. *American journal of respiratory and critical care medicine* 203: 977-986, 2021.

7. **Ramsey KA, Rosenow T, Turkovic L, Skoric B, Banton G, Adams AM, Simpson SJ, Murray C, Ranganathan SC, Stick SM, and Hall GL**. Lung Clearance

Index and Structural Lung Disease on Computed Tomography in Early Cystic Fibrosis. *American journal of respiratory and critical care medicine* 193: 60-67, 2016.

8. **Kent L, Reix P, Innes JA, Zielen S, Le Bourgeois M, Braggion C, Lever S, Arets HG, Brownlee K, Bradley JM, Bayfield K, O'Neill K, Savi D, Bilton D, Lindblad A, Davies JC, Sermet I, and De Boeck K**. Lung clearance index: evidence for use in clinical trials in cystic fibrosis. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 13: 123-138, 2014.

9. **Saunders C, Jensen R, Robinson PD, Stanojevic S, Klingel M, Short C, Davies JC, and Ratjen F**. Integrating the multiple breath washout test into international multicentre trials. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 19: 602-607, 2020.

10. **Subbarao P, Milla C, Aurora P, Davies JC, Davis SD, Hall GL, Heltshe S, Latzin P, Lindblad A, Pittman JE, Robinson PD, Rosenfeld M, Singer F, Starner TD, Ratjen F, and Morgan W**. Multiple-Breath Washout as a Lung Function Test in Cystic Fibrosis. A Cystic Fibrosis Foundation Workshop Report. *Annals of the American Thoracic Society* 12: 932-939, 2015.

11. **Ratjen F, Davis SD, Stanojevic S, Kronmal RA, Hinckley Stukovsky KD, Jorgensen N, and Rosenfeld M**. Inhaled hypertonic saline in preschool children with cystic fibrosis (SHIP): a multicentre, randomised, double-blind, placebo-controlled trial. *The Lancet Respiratory medicine* 7: 802-809, 2019.

12. **Ratjen F, Hug C, Marigowda G, Tian S, Huang X, Stanojevic S, Milla CE, Robinson PD, Waltz D, and Davies JC**. Efficacy and safety of lumacaftor and ivacaftor in patients aged 6-11 years with cystic fibrosis homozygous for F508del-CFTR: a randomised, placebo-controlled phase 3 trial. *The Lancet Respiratory medicine* 5: 557-567, 2017.

13. **Davies JC, Sermet-Gaudelus I, Naehrlich L, Harris RS, Campbell D, Ahluwalia N, Short C, Haseltine E, Panorchan P, Saunders C, Owen CA, and Wainwright CE**. A phase 3, double-blind, parallel-group study to evaluate the efficacy and safety of tezacaftor in combination with ivacaftor in participants 6 through 11 years of age with cystic fibrosis homozygous for F508del or heterozygous for the F508del-CFTR mutation and a residual function mutation. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 20: 68-77, 2021.

14. **Nielsen JG**. Lung clearance index: should we really go back to nitrogen washout? *The European respiratory journal* 43: 655-656, 2014.

15. **Kane M, Rayment JH, Jensen R, McDonald R, Stanojevic S, and Ratjen F**. Correcting for tissue nitrogen excretion in multiple breath washout measurements. *PloS one* 12: e0185553, 2017.

16. **Guglani L, Kasi A, Starks M, Pedersen KE, Nielsen JG, and Weiner DJ**. Difference between SF(6) and N(2) Multiple Breath Washout kinetics is due to N(2) back diffusion and error in N(2) offset. *Journal of applied physiology (Bethesda, Md : 1985)* 2018.

17. **Cummings B, Hamilton ML, Ciaffoni L, Pragnell TR, Peverall R, Ritchie GA, Hancock G, and Robbins PA**. Laser-based absorption spectroscopy as a technique for rapid in-line analysis of respired gas concentrations of O2 and CO2. *Journal of applied physiology (Bethesda, Md : 1985)* 111: 303-307, 2011.

18. **Arieli R, Ertracht O, and Daskalovic Y**. Infrared CO2 analyzer error: an effect of background gas (N2 and O2). *Journal of applied physiology (Bethesda, Md : 1985)* 86: 647-650, 1999.

19. **Oxigraf Inc**. Technology - Laser Absorption Spectroscopy Summary, *retrieved 5th of July 2021 from https://www.oxigraf.com/technology/*

20. **Sandvik RM, Gustafsson PM, Lindblad A, Robinson PD, and Nielsen KG**. Improved agreement between N(2) and SF(6) multiple breath washout in healthy infants and toddlers with improved EXHALYZER D(®) sensor performance. *Journal of applied physiology (Bethesda, Md : 1985)* 2021.

21. **Anagnostopoulou P, Latzin P, Jensen R, Stahl M, Harper A, Yammine S, Usemann J, Foong RE, Spycher B, Hall GL, Singer F, Stanojevic S, Mall MA, Ratjen F, and Ramsey KA**. Normative data for multiple breath washout outcomes in school-aged Caucasian children. *The European respiratory journal* 55: 1901302, 2020.

22. **Korten I, Kieninger E, Yammine S, Regamey N, Nyilas S, Ramsey K, Casaulta C, Latzin P, and For The Scild Study G**. The Swiss Cystic Fibrosis Infant Lung Development (SCILD) cohort. *Swiss medical weekly* 148: w14618, 2018.

23. **Bayfield KJ, Horsley A, Alton E, Irving S, Bush A, and Davies JC**. Simultaneous sulfur hexafluoride and nitrogen multiple-breath washout (MBW) to examine inherent differences in MBW outcomes. *ERJ open research* 5: 00234-02018, 2019.

24. **Bell AS, Lawrence PJ, Singh D, and Horsley A**. Feasibility and challenges of using multiple breath washout in COPD. *International journal of chronic obstructive pulmonary disease* 13: 2113-2119, 2018.

25. **Jensen R, Stanojevic S, Gibney K, Salazar JG, Gustafsson P, Subbarao P, and Ratjen F**. Multiple breath nitrogen washout: a feasible alternative to mass spectrometry. *PloS one* 8: e56868, 2013.

26. **Singer F, Houltz B, Latzin P, Robinson P, and Gustafsson P**. A realistic validation study of a new nitrogen multiple-breath washout system. *PloS one* 7: e36083, 2012.

27. **Robinson PD, Latzin P, Verbanck S, Hall GL, Horsley A, Gappa M, Thamrin C, Arets HG, Aurora P, Fuchs SI, King GG, Lum S, Macleod K, Paiva M, Pillow JJ, Ranganathan S, Ratjen F, Singer F, Sonnappa S, Stocks J, Subbarao P, Thompson BR, and Gustafsson PM**. Consensus statement for inert gas washout measurement using multiple- and single- breath tests. *The European respiratory journal* 41: 507-522, 2013.

28. **Wyler F, Oestreich M-A, Frauchiger BS, Ramsey K, and Latzin P**. Correction of measurement error in a commercial multiple-breath washout device. *medRxiv* [preprint] 2021.2002.2006.21251250, 2021.

29. **Poncin W, Singer F, Aubriot AS, and Lebecque P**. Agreement between multiple-breath nitrogen washout systems in children and adults. *Journal of cystic fibrosis : official journal of the European Cystic Fibrosis Society* 16: 258-266, 2017.

30. **Yammine S, Summermatter S, Singer F, Lauener R, and Latzin P**. Feasibility of nitrogen multiple-breath washout in inexperienced children younger than 7 years. *Pediatric pulmonology* 51: 1183-1190, 2016.

31. **Yammine S, Singer F, Abbas C, Roos M, and Latzin P**. Multiple-breath washout measurements can be significantly shortened in children. *Thorax* 68: 586-587, 2013.

32. **Foong RE, Harper AJ, Skoric B, King L, Turkovic L, Davis M, Clem CC, Rosenow T, Davis SD, Ranganathan S, Hall GL, and Ramsey KA**. The clinical utility of lung clearance index in early cystic fibrosis lung disease is not impacted by the number of multiple-breath washout trials. *ERJ open research* 4: 00094-02017, 2018.

33. **Frauchiger BS, Carlens J, Herger A, Moeller A, Latzin P, and Ramsey KA**. Multiple breath washout quality control in the clinical setting. *Pediatric pulmonology* 56: 105-112, 2021.

34. **Nielsen N, Nielsen JG, and Horsley AR**. Evaluation of the impact of alveolar nitrogen excretion on indices derived from multiple breath nitrogen washout. *PloS one* 8: e73335, 2013.

35. **Sandhu D, Ritchie GAD, and Robbins PA**. The differing physiology of nitrogen and tracer gas multiple-breath washout techniques. *ERJ open research* 7: 2021.

## FIGURES

**Figure 1:** $N_2$ error as a function of $N_2$ in the gas mixtures. The $N_2$ error here is the absolute difference between measured $N_2$ and reference $N_2$ as a function of reference $N_2$ in the gas mixtures. Dashed curves represent the combined fits through the errors of the individual gas sensor errors, for selected concentrations of $CO_2$. Dotted vertical line indicates the end of test condition. The color shading indicates the reference $CO_2$ concentrations of the gas mixtures. Dots represent mean of 6 measurements (triplicates on 2 devices) performed with 12 technical gas mixtures as reference ($CO_2$: 0%, 2.5%, 7.5%, $O_2$: 30%, 60%, 90%, Rest, $N_2$: Rest), triangles represent mean of 3 mass spectrometry reference measurements of 7 mixtures at the end of test condition ($N_2$: 2%, $CO_2$: 0%, 1%, 2%, 3%, 4%, 5%, 6%, $O_2$: Rest). Error bars represent SD of measurements for each mixture. For an overview of the gas mixtures see OLS (Technical gases and Table 2).

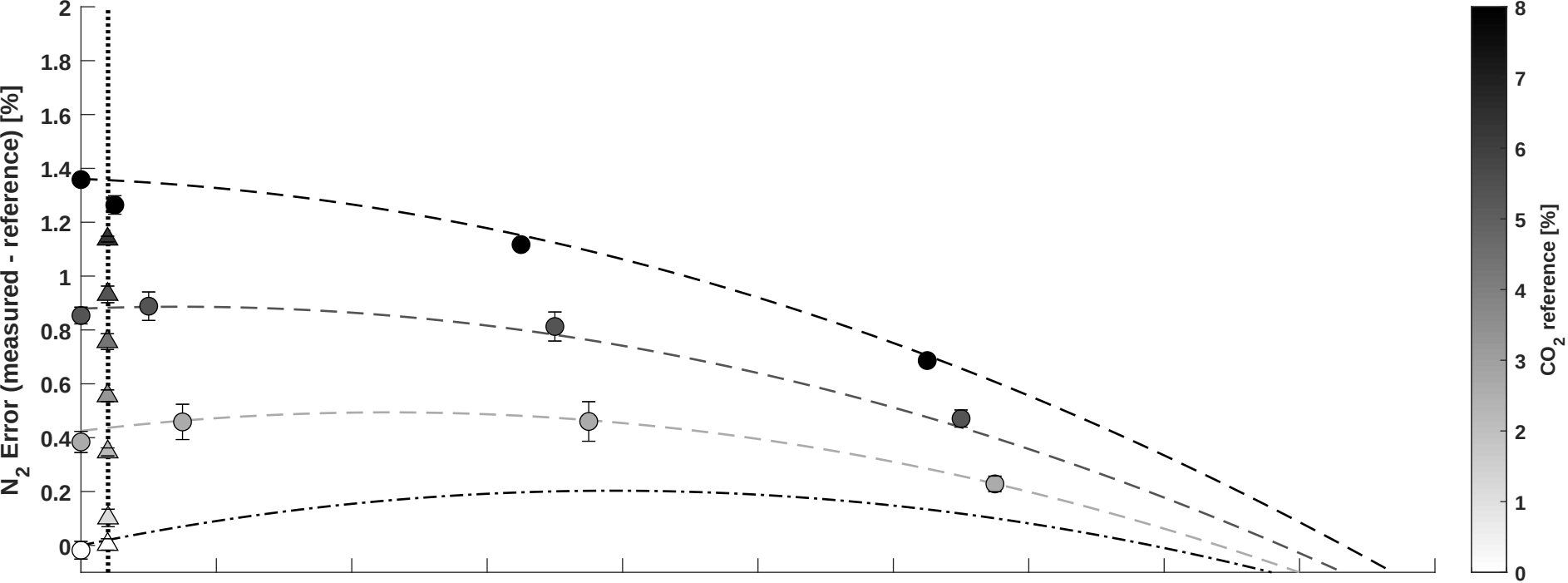
572 **Figure 2**: Illustration of the effect of the sensor correction on the $N_2$ signal and
573 consequently on the end of test in an example MBW measurement. Traced in gray is
574 the signal output of the standard signal processing, the corrected signal is shown in
575 black. Vertical dashed lines represent the end of test for the original standard and
576 corrected measurement respectively. The dashed horizontal line corresponds to
577 $1/40^{th}$ of the initial $N_2$ concentration (end of test, ca. 2% $N_2$). Dashed vertical lines
578 represent the original end of test (end of test condition reached in standard
579 processing) and new end of test (end of test condition reached in corrected
580 processing). **(A)** Time course of $N_2$ throughout a standard MBW measurement. **(B)**
581 Zoom into the critical period of end of test determination. In this example the test
582 ends 5 breaths earlier in the corrected measurement compared to standard.

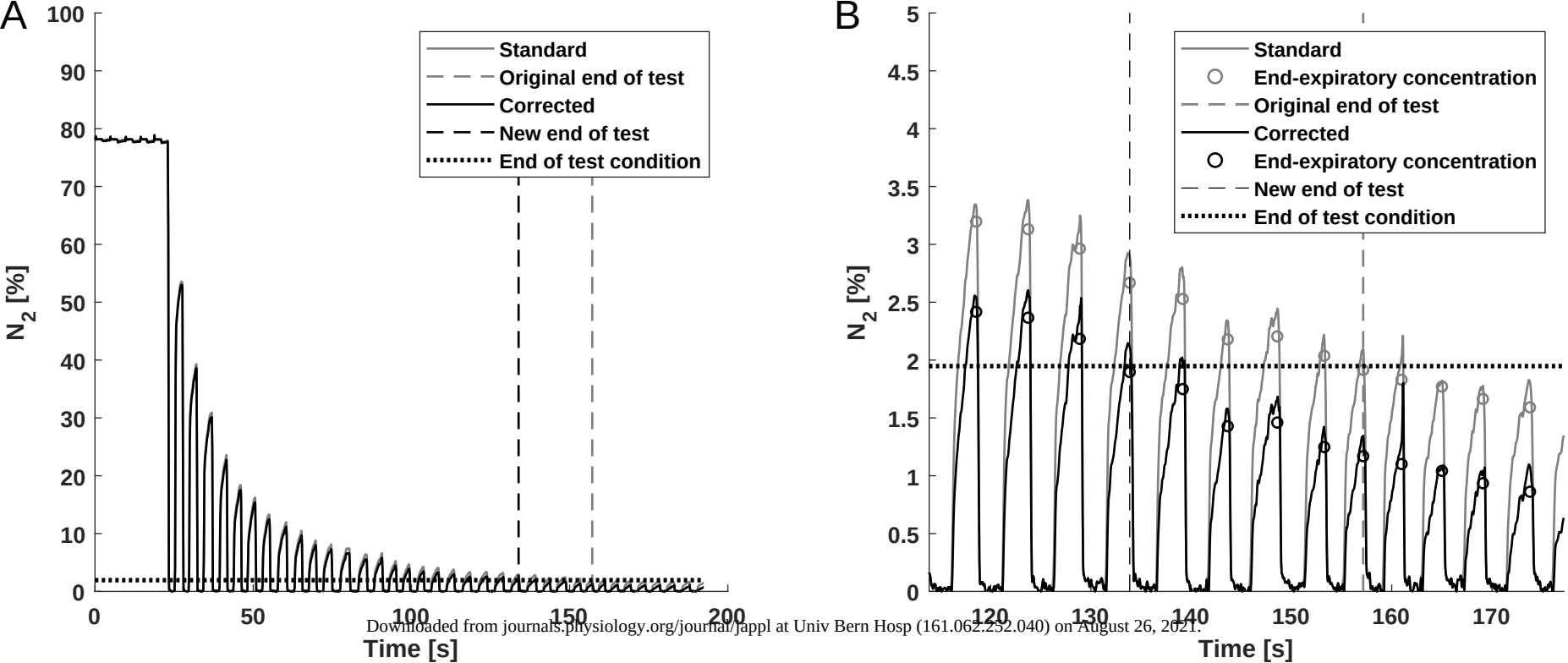
**Figure 3:** Observed absolute error between reference and measured gas concentrations. (dots: mean of error of one gas mixture, error bars: +/- SD of error). Curves represent a two parameter quadratic polynomial fitted through the error values (see OLS for details), represented here as dashed curves for given $CO_2$ concentrations. Dots represent mean of 6 measurements (triplicates on 2 devices) performed with 12 technical gas mixtures as reference ($CO_2$: 0%, 2.5%, 7.5%, $O_2$: 30%, 60%, 90%, Rest, $N_2$: Rest), triangles represent mean of 3 mass spectrometry reference measurements of 7 mixtures at the end of test condition ($N_2$: 2%, $CO_2$: 0%, 1%, 2%, 3%, 4%, 5%, 6%, $O_2$: Rest). Error bars represent SD of measurements for each mixture. For an overview of the gas mixtures see OLS (Technical gases and Table 2). (A) Absolute $O_2$ error as a function of $O_2$ and $CO_2$ concentration, (B) Absolute $CO_2$ error as a function of $O_2$ and $CO_2$ concentration.

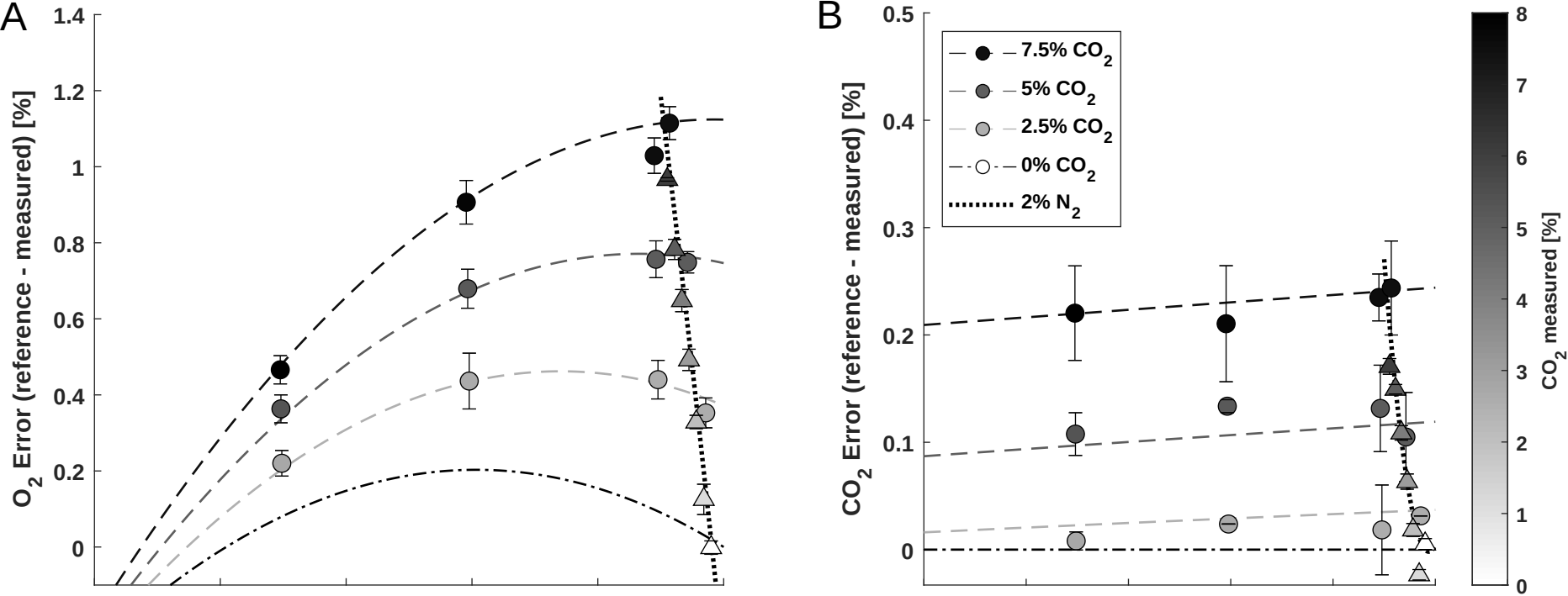
**Figure 4**: Multiple-breath washout outcomes **(A)** Lung Clearance Index (LCI) and **(B)** functional residual capacity (FRC) after sensor correction (corrected) vs standard (standard; Spiroware 3.3) in healthy controls (HC) and patients with cystic fibrosis (CF). Solid black line indicates line of equality, dashed line represents a linear fit through the data points.

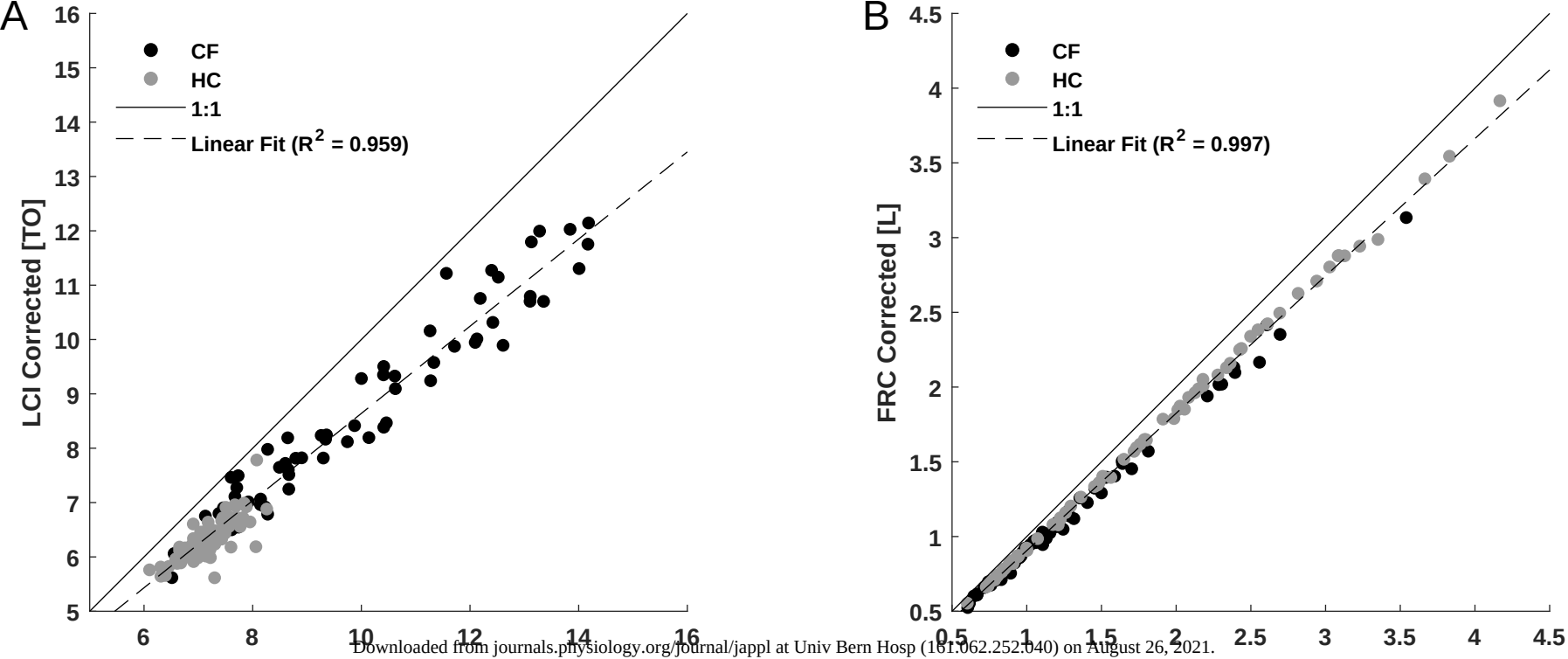
**Figure 5**: Bland-Altman plot of the absolute difference (corrected – standard) of multiple-breath washout outcomes **(A)** Lung Clearance Index (LCI) in turnover [TO] and **(B)** functional residual capacity (FRC) in liter [L] of healthy controls (HC) and patients with cystic fibrosis (CF) due to sensor correction, plotted against the mean outcomes (mean of corrected and standard).

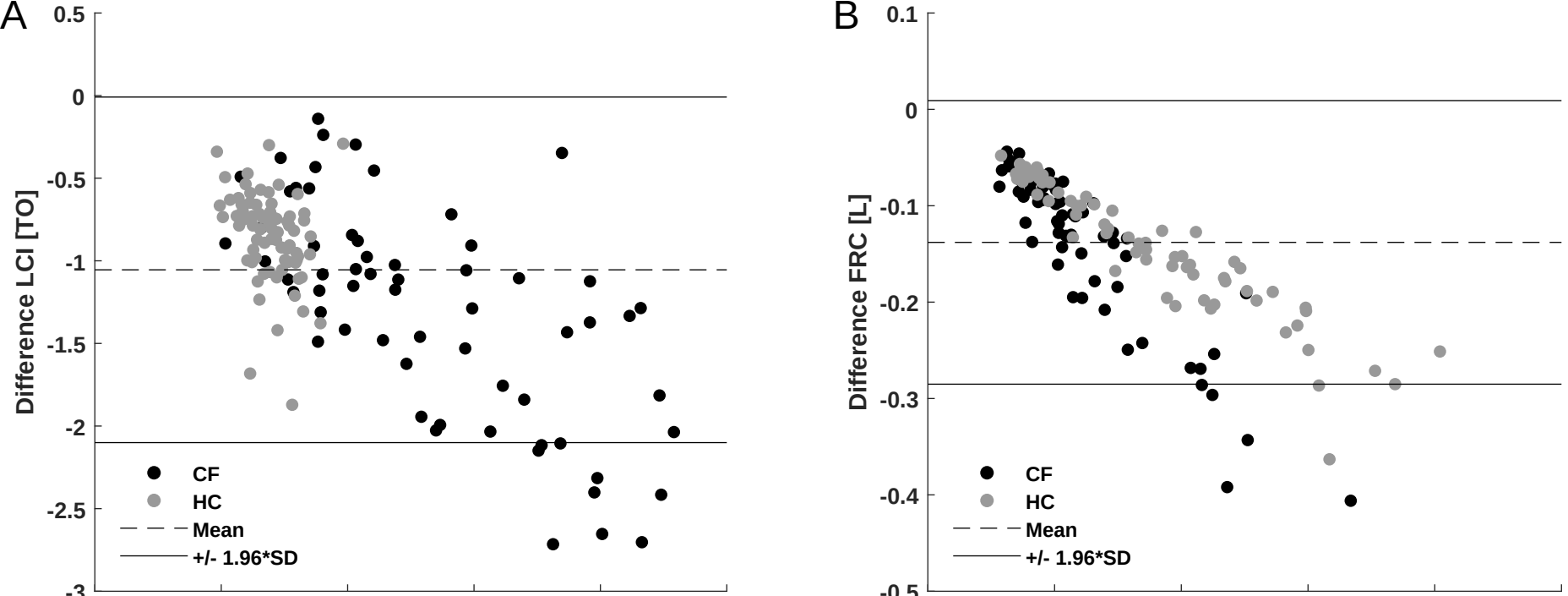
**Figure 6**: Illustration of the effect of the sensor correction function on nitrogen measurement in the late phase of MBW tests. **(A)** Example of the equlibrium $N_2$ reached in a very long continued MBW measurement, displaying a greatly decreased $N_2$-back-diffusion equilibrium (tissue nitrogen). **(B)** Corrected $N_2$ plotted against standard $N_2$ in conditions around the end of test condition (2% $N_2$).











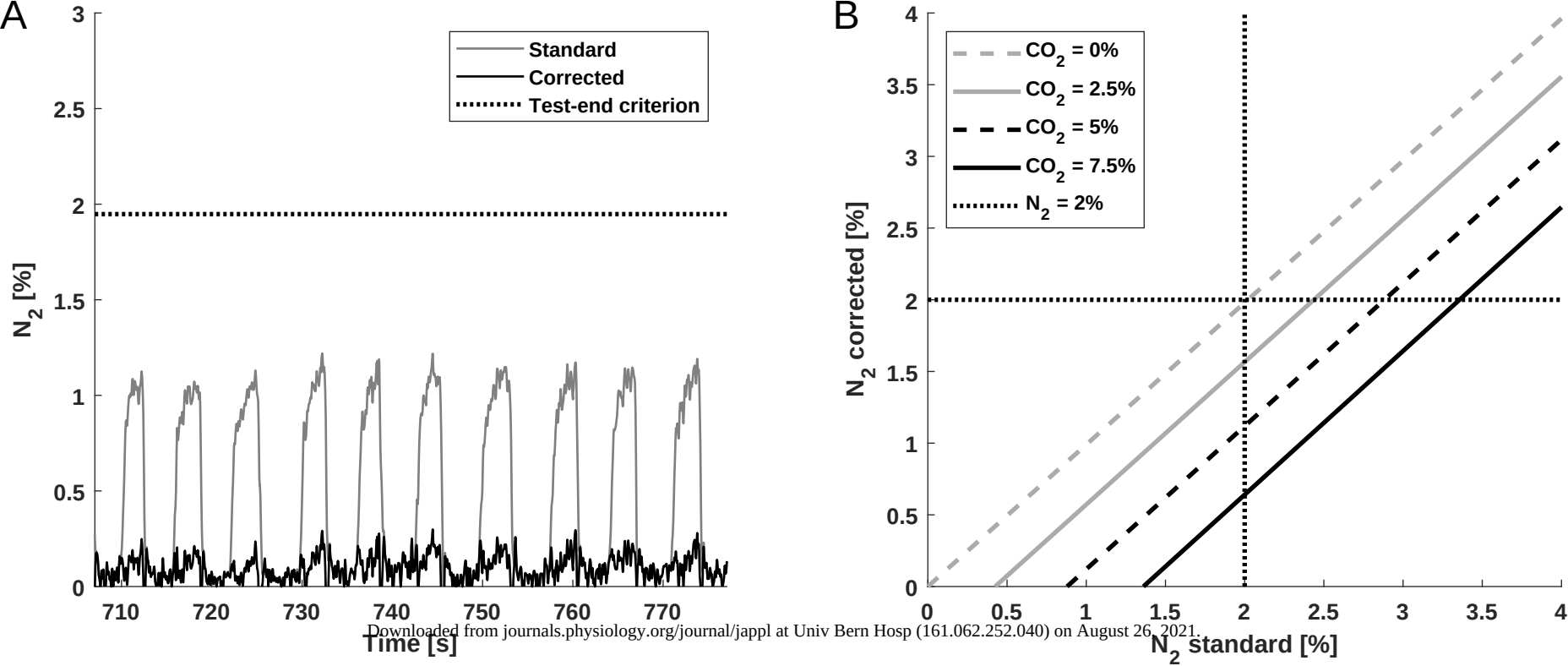


**Table 1**

| Condition | Signal Processing | | | | | | $N_2$ Error | | Contribution | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Standard | | | Corrected | | | | | | |
| | $[N_2]$ | $[CO_2]$ | $[O_2]$ | $[N_2]$ | $[CO_2]$ | $[O_2]$ | abs | rel | $[CO_2]$ | $[O_2]$ |
| Original end of test [%] | 2.00 | 5.00 | 93.0 | 1.10 | 5.12 | 93.8 | 0.90 | 82.4 | 13.5 | 86.5 |
| New end of test [%] | 2.88 | 4.88 | 92.2 | 2.00 | 5.00 | 93.0 | 0.88 | 44.1 | 13.3 | 86.7 |
| No nitrogen [%] | 0.88 | 4.88 | 94.2 | 0.00 | 5.00 | 95.0 | 0.88 | - | - | - |

**Table 1**: Specific examples of sensor impact on measurement of $N_2$ in three conditions of interest. The original end of test corresponds to a gas mixture that would be identified as the end of test in standard processing. The second condition corresponds to the new end of test after sensor correction. The third condition contains no real nitrogen. Standard concentrations denote concentrations measured in standard Spiroware 3.3 processing. Corrected concentrations correspond to concentrations after sensor correction is applied. $N_2$ error summarizes the absolute (abs) difference between $N_2$ in standard vs. corrected, as well as the relative (rel) error ((standard-corrected)/corrected). The relative contribution of each sensor in [%] to the total error in N2 concentration is listed under "Contribution".

## Table 2

| | | | Standard | | | Corrected | | | Difference | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | n | mean | SD | p-value* | mean | SD | p-value* | mean | rel [%] | 95% CI [%] | | | p-value† |
| **LCI** [TO] | All | 147 | 8.33 | 2.05 | | 7.31 | 1.7 | | 1.02 | 11.9 | 11.2 | - | 12.5 | **<0.001** |
| | HC | 85 | 7.12 | 0.51 | | 6.30 | 0.4 | | 0.82 | 11.3 | 10.6 | - | 12.1 | **<0.001** |
| | CF | 62 | 9.99 | 2.21 | | 8.69 | 1.8 | | 1.30 | 12.6 | 11.4 | - | 13.8 | **<0.001** |
| | Difference | | -2.87 | | **<0.001** | -2.38 | | **<0.001** | | | | | | |
| **FRC** [L] | All | 147 | 1.63 | 0.87 | | 1.49 | 0.80 | | 0.14 | 8.9 | 8.6 | - | 9.3 | **<0.001** |
| | HC | 85 | 1.87 | 0.95 | | 1.73 | 0.89 | | 0.14 | 7.9 | 7.6 | - | 8.1 | **<0.001** |
| | CF | 62 | 1.31 | 0.61 | | 1.17 | 0.53 | | 0.14 | 10.4 | 9.7 | - | 11.0 | **<0.001** |
| | Difference | | 0.56 | | **<0.001** | 0.56 | | **<0.001** | | | | | | |
| **CEV** [L] | All | 147 | 14.9 | 7.4 | | 11.9 | 5.7 | | 3.03 | 19.6 | 18.8 | - | 20.4 | **<0.001** |
| | HC | 85 | 14.6 | 6.7 | | 12.0 | 5.6 | | 2.67 | 18.2 | 17.4 | - | 19.0 | **<0.001** |
| | CF | 62 | 15.2 | 8.3 | | 11.7 | 5.8 | | 3.54 | 21.5 | 20.1 | - | 23.0 | **<0.001** |
| | Difference | | -0.58 | | 0.6532 | 0.29 | | 0.7601 | | | | | | |

**Table 2**: Summary of the differences in Lung Clearance Index (LCI), functional residual capacity (FRC) and cumulative expired volume (CEV) between healthy controls (HC) and and patients with cystic fibrosis (CF) in the retrospective dataset before (standard) and after (corrected) the application of the sensor correction function. *unpaired t test; †paired t test. Bold print indicates statistical significance.